UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/539,450	12/23/2005	Hisashi Narimatsu	159-90	6812

23117 7590 11/17/2006

NIXON & VANDERHYE, PC
901 NORTH GLEBE ROAD, 11TH FLOOR
ARLINGTON, VA 22203

EXAMINER	
RAGHU, GANAPATHIRAM	
ART UNIT	**PAPER NUMBER**
1652	

DATE MAILED: 11/17/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/539,450	NARIMATSU ET AL.	
	Examiner	Art Unit	
	Ganapathirama Raghu	1652	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *22 September 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-20* is/are pending in the application.

4a) Of the above claim(s) *7-20* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-6* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☒ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *07/05,04/06, 05/06*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☒ Other: *SEQ ALIGN*.

DETAILED ACTION

Claims 1-20 are pending in this application and claims 1-6 are now under consideration for examination. Claims 7-20 are withdrawn as they are drawn to non-elected inventions.

Election/Restrictions

Applicants' election with traverse of Group I, claims 1-6 and SEQ ID NOs: 2 and 4 for prosecution in their response dated 22 Sep. 2006 is acknowledged. The traversal is on the grounds there would not be serious burden on the examiner to examine groups I and II (polynucleotide group), therefore restriction between groups be withdrawn and applicants' have for examination of all the claims pertaining to groups I and II and furthermore polypeptide sequences of SEQ ID NO: 2 and SEQ ID NO: 4 are related in structure and function. Applicants' arguments have been considered, examiner agrees with the arguments regarding the structure and function relationship of SEQ ID NOs.: 2 and 4 and therefore restriction requirement between them has been withdrawn, however, respectfully disagrees with the argument that searching all claims is "not a serious burden" for the following reasons. Searching structurally distinct molecules like the polypeptides of group III (antibody group) and the polynucleotides of group II are not coextensive and involves search of different databases and non-patent literature, as prior to the concomitant isolation and expression of the sequence of interest there may be scientific journal articles devoted solely to the polypeptides which would not have described the polynucleotide and moreover the polypeptides may have been isolated by biochemical means. Group I polypeptides encompasses molecules which are structurally distinct and claimed in terms of variants with a wide ranging percentage sequence identity and amino acid changes to

SEQ ID NO: 2 or SEQ ID NO: 4 and therefore encoding polynucleotides encompassed by these claims are very broad and thus a combined search involving polypeptides and encoding polynucleotides of the instant invention and analysis of results would be a serious search burden. Therefore, for the above-cited reasons and contrary to applicants' argument, the requirement is still deemed proper and is therefore made FINAL.

Priority

Acknowledgment is made of applicant's claim for foreign priority under 35 U.S.C. 119(a)-(d). This application is a 371 of PCT/JP04/00608 filed on 01/23/2004 and claims the priority date of Japanese application 2003-014792, 2003-285310 and 2003-392555 filed on 01/23/2003, 08/01/2003 and 11/21/2003 respectively. However, Examiner notes that the English translation for the said applications are not provided.

Information Disclosure Statement

The information disclosure statement (IDS) submitted on 07/06/ 2005, 04/26/2006 and 05/25/2006 are in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

Drawings

The drawings are considered for examination purposes only.

Claim Rejections 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-6 are rejected under 35 U.S.C. 101 because the claims could read on a non-statutory subject matter. The claims are drawn to an "A β1,3-N-acetyl-D-galactosamine transferase", which could read on product of nature. Claims directed to such matter are considered non-statutory. Examiner suggests amending the claims to recite 'An isolated β1,3-N-acetyl-D-galactosamine transferase'' to show the hand of man.

Claim Rejections: 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claim 2 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Claim 2 recites the phrase "... pH range of 6.2 to 6.6 than in other pH ranges", the metes and bounds of the phrase is not clear, does this include any pH below 6.2 or any pH above 6.6?. Clarification and correction is required. Furthermore, clearly this cannot in fact be the case in view of proteins instability in strong acids and bases, clearly no enzyme will have higher activity at pH 0-1 and 13-14 than at pH 6.2-6.6.

Claim 6 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the

invention. Claim 16 recites the phrase "... 30% identity with an amino acid sequence covering amino acids 189 to 500 shown in SEQ ID NO: 2 or 35 to 504 shown in SEQ ID NO: 4", the metes and bounds of the phrase is not clear and the examiner suggests changing the phrase to "... 30% sequence identity with an amino acid sequence covering amino acids 189 to 500 shown in SEQ ID NO: 2 or 35 to 504 shown in SEQ ID NO: 4. Correction is required.

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 1-3 are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not disclosed in the specification in such a way as to reasonably convey to one of skilled in the relevant art that the invention(s), at the time the application was filed, had possession of the claimed invention.

Claims 1-3 are directed to a genus of polypeptides having β1,3-N-acetyl-D-galactosamine transferase activity. The specification does not contain any disclosure of the structure of all the polypeptide sequences included in the claimed genera. The genus of polypeptides claimed is large variable genus with the potentiality of many different structures. Therefore, many structurally distinct polypeptides are encompassed within the scope of the claims. The specification discloses only two sequences of claimed genus (i. e. that of SEQ ID NO: 2 or 4), which is insufficient to put one of skill in the art in possession of the attributes and features of all species within the claimed genus. The structure of any polypeptide having the β1,3-N-acetyl-D-galactosamine transferase activity is completely undefined and the specification

does not define the structural features necessary for members of the genus to be selected. Therefore, one skilled in the art cannot reasonably conclude that the applicant had possession of the claimed invention at the time the instant application was filed. Applicant is referred to the revised guidelines concerning compliance with the written description requirement of U.S.C. 112, first paragraph, published in the Official Gazette and also available at www.uspto.gov.

Claim 6 is rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

Claim 6 is directed to an isolated β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2, or SEQ ID NO: 4 or variants of said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical characteristics. Claim 6 is rejected under this section 35 U.S.C. 112, because the claims are directed to a genus of polypeptides with no support in the specification for the structural details associated with the function i.e., an isolated β1,3-N-acetyl-D- galactosamine transferase polypeptide having specific activity and biochemical characteristics. No description of identifying characteristics of all of the sequences of an isolated β1,3-N-acetyl-D- galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or variants of

said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical characteristics has been provided by the applicants in the specification. No information, beyond the characterization of the β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 has been provided by the applicants, which would indicate that they had possession of the claimed genus of the polypeptides i.e., an isolated β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or variants of said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical characteristics, as 30% sequence identity corresponds to a large variation in the structure and structures with such a large variation may not have similar functional characteristics in terms substrate specificity or kinetic/catalytic properties. Therefore, one skilled in the art cannot reasonably conclude that applicant had possession of the claimed invention at the time the instant application was filed. Applicant is referred to the revised guidelines concerning compliance with the written description requirement of U.S.C. 112, first paragraph, published in the Official Gazette and also available at www.uspto.gov.

Claims 1-3 and 6 are rejected under 35 U.S.C. 112, first paragraph, because the specification, while being enabling for an isolated β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising the amino

acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, does not reasonably provide enablement for any isolated β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or variants of said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical characteristics. The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and or use the invention commensurate in scope with the claims.

Factors to be considered in determining whether undue experimentation is required are summarized in *In re Wands* (858 F.2d 731, 8 USPQ 2nd 1400 (Fed. Cir. 1988)) as follows: (1) the quantity of experimentation necessary, (2) the amount of direction or guidance presented, (3) the presence or absence of working examples, (4) the nature of the invention, (5) the state of the prior art, (6) the relative skill of those in the art, (7) the predictability or unpredictability of the art, and (8) the breadth of the claim(s).

Claims 1-3 and 6 are so broad as to encompass for any isolated β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or variants of said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical characteristics. The scope of the claims are not commensurate with the enablement provided by the disclosure with regard to the extremely large number of polypeptides and encoding polynucleotides broadly encompassed by the claims. Since the amino acid sequence of

a protein encoded by a polynucleotide determines its structural and functional properties, predictability of which changes can be tolerated in a protein's amino acid sequence and obtain the desired activity requires knowledge and guidance with regard to which amino acids in the protein's sequence and the respective codons in its polynucleotide, if any, are tolerant of modification and which are conserved (i.e. expectedly intolerant to modification), and detailed knowledge of the ways in which the encoded proteins' structure relates to its function. However, in this case the disclosure is limited to an isolated β1,3-N-acetyl-D- galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, but provides no guidance with regard to the making of variants and mutants or with regard to other uses. In view of the great breadth of the claims, amount of experimentation required to make the claimed polypeptides and encoding polynucleotides, the lack of guidance, working examples, and unpredictability of the art in predicting function from a polypeptide primary structure (e.g., see Ngo et al. in The Protein Folding Problem and Tertiary Structure Prediction, 1994, Merz et al. (ed.), Birkhauser, Boston, MA, pp. 433 and 492-495), the claimed invention would require undue experimentation. As such, the specification fails to teach one of ordinary skill how to use the full scope of the polypeptides encompassed by this claim.

While enzyme isolation techniques, recombinant and mutagenesis techniques are known, and it is not routine in the art to screen for multiple substitutions or multiple modifications as encompassed by the instant claim, the specific amino acid positions within a protein's sequence where amino acid modifications can be made with a reasonable expectation of success in obtaining the desired activity/utility are limited in any protein and the result of such

modifications is unpredictable. In addition, one skilled in the art would expect any tolerance to modification for a given protein to diminish with each further and additional modification, e.g. multiple substitutions or deletions.

The specification does not support the broad scope of the claims which encompass all modifications to any isolated β1,3-N-acetyl-D- galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or variants of said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical characteristics, because the specification does not establish: (A) regions of the protein/polynucleotide structure which may be modified without affecting the activity of encoded β1,3-N-acetyl-D- galactosamine transferase polypeptide having specific activity and biochemical characteristics; (B) the general tolerance of the polypeptide and the polynucleotide encoding β1,3-N-acetyl-D- galactosamine transferase polypeptide having specific activity and biochemical characteristics to modification and extent of such tolerance; (C) a rational and predictable scheme for modifying any amino acid residue or the respective codon in the polynucleotide with an expectation of obtaining the desired biological function; and (D) the specification provides insufficient guidance as to which of the essentially infinite possible choices is likely to be successful.

Thus, applicants have not provided sufficient guidance to enable one of ordinary skill in the art to make and use the claimed invention in a manner reasonably correlated with the scope of the claim broadly including polynucleotides with an enormous number of modifications. The scope of the claim must bear a reasonable correlation with the scope of enablement (*In re Fisher*,

166 USPQ 19 24 (CCPA 1970)). Without sufficient guidance, determination of any isolated β1,3-N-acetyl-D- galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 40%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or variants of said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical/biological characteristics is unpredictable and the experimentation left to those skilled in the art is unnecessarily, and improperly, extensive and undue. See *In re Wands* 858 F.2d 731, 8 USPQ2nd 1400 (Fed. Cir, 1988).

Claim Rejections 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1-6 are rejected under 35 U.S.C. 102(b) as being anticipated by Strausberg et al., (PNAS., 2002, Vol. 99 (26): 16899-16903). Claims 1-6 are directed to any isolated β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or to a polypeptide having an amino acid sequence covering amino acids 189 to 500 of SEQ ID NO: 2 or to a polypeptide having an amino acid sequence covering amino acids 36 to 500 of SEQ ID NO: 2 or SEQ ID NO: 4 or to a polypeptide having an amino acid sequence covering amino acids 35 to 504 of SEQ ID NO: 4 or variants of

said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific activity and biochemical characteristics. Strausberg et al., (*supra*) teach the isolation of a polypeptide (B3 GALNT2; ORF Name= RP4-534P7.1-001) annotated as β1,3-N-acetyl-D- galactosamine transferase that has 100% sequence homology to SEQ ID NO: 2 or to a polypeptide having an amino acid sequence covering amino acids 189 to 500 of SEQ ID NO: 2 or to a polypeptide having an amino acid sequence covering amino acids 36 to 500 of SEQ ID NO: 2 of the instant application (see sequence alignment provided). The reference is silent regarding the specific activity and biochemical characteristics of said polypeptide, however examiner takes the position that said polypeptide by virtue of 100% sequence homology to SEQ ID NO: 2 also inherently posses the same specific activity and biochemical characteristics as that of SEQ ID NO: 2. The reference also teaches encoding polynucleotides, vectors, host cells and method of making the polypeptide. Therefore, Strausberg et al., anticipate claims 1-6 as written.

Claims 1-3 and 6 are rejected under 35 U.S.C. 102(b) as being anticipated by Kawai et al., (Nature, 2001, Vol. 409: 685-690). Claims 1-3 and 6 are directed to any isolated β1,3-N-acetyl-D-galactosamine transferase polypeptide having specific activity and biochemical characteristics and comprising an amino acid sequence having 30%-99% sequence identity to the amino acid sequence of SEQ ID NO: 2 or to a polypeptide having an amino acid sequence covering amino acids 189 to 500 of SEQ ID NO: 2 or to a polypeptide having an amino acid sequence covering amino acids 36 to 500 of SEQ ID NO: 2 or SEQ ID NO: 4 or to a polypeptide having an amino acid sequence covering amino acids 35 to 504 of SEQ ID NO: 4. or variants of said sequences in which one or several amino acids are substituted, deleted or inserted and having said specific

activity and biochemical characteristics. Kawai et al., (*supra*) teach the isolation of a polypeptide (for clone identity http://genomec.gsc.riken.go.jp/genome/fantom1/fig./nature/supplement/;user name: fantom1; password:fntm0828) annotated as β1,3-N-acetyl-D- galactosamine transferase that has 100% sequence homology to SEQ ID NO: 4 or to a polypeptide having an amino acid sequence covering amino acids 35 to 504 of SEQ ID NO: 4 of the instant application (see sequence alignment provided). The reference is silent regarding the specific activity and biochemical characteristics of said polypeptide, however examiner takes the position that said polypeptide by virtue of 100% sequence homology to SEQ ID NO: 4 also inherently posses the same specific activity and biochemical characteristics as that of SEQ ID NO: 4. The reference also teaches encoding polynucleotides, vectors, host cells and method of making the polypeptide. Therefore, Kawai et al., anticipate claims 1-3 and 6 as written.

Conclusion

None of the claims are allowed.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ganapathirama Raghu whose telephone number is 571-272-4533. The examiner can normally be reached on 8 am - 4.30 pm. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Ponnathapu Achutamurthy can be reached on 571-272-0928. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300 for regular communications and for After Final communications.

Any inquiry of a general nature or relating to the status of the application or proceeding should be directed to the receptionist whose telephone number is 571-272-1600. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Ganapathirama Raghu, Ph.D.
Patent Examiner
Art Unit 1652

Oct. 22, 2006.

REBECCA E. PROUTY
PRIMARY EXAMINER
GROUP ~~1800~~ 1600

GenCore version 5.1.9

OM protein - protein search, using sw model

Run on: October 3, 2006, 10:00:07 ; Search time 142.625 Seconds
(without alignments)
3268.776 Million cell updates/sec

Title: US-10-539-450-4
Perfect score: 2707
Sequence: 1 MRNWLVLLCPCVLGAALHLW.....KLWELKELCGDPCQCEAKVR 504

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2707	100.0	504	2	Q8BG28_MOUSE	Q8bg28 m adult mal
2	2698	99.7	504	2	Q5U4F9_MOUSE	Q5u4f9 mus musculu
3	2369	87.5	500	2	Q8NCR0_HUMAN	Q8ncr0 homo sapien
4	2068.5	76.4	427	2	Q59GR3_HUMAN	Q59gr3 homo sapien
5	1823	67.3	488	2	Q5M900_XENTR	Q5m900 xenopus tro
6	1821	67.3	486	2	Q6NRQ1_XENLA	Q6nrq1 xenopus lae
7	1338.5	49.4	491	2	Q502B3_BRARE	Q502b3 brachydanio
8	1261.5	46.6	326	2	Q5TCI3_HUMAN	Q5tci3 homo sapien
9	1261.5	46.6	382	2	Q96AL7_HUMAN	Q96al7 homo sapien
10	1096.5	40.5	426	2	Q4RPY6_TETNG	Q4rpy6 tetraodon n
11	997	36.8	209	2	Q8BXL0_MOUSE	Q8bxl0 mus musculu
12	606	22.4	155	2	Q4TDH0_TETNG	Q4tdh0 tetraodon n
13	255.5	9.4	663	2	Q6ZKZ3_ORYSA	Q6zkz3 oryza sativ
14	246	9.1	673	2	Q8GXG6_ARATH	Q8gxg6 arabidopsis
15	243.5	9.0	343	2	Q5BL85_XENTR	Q5bl85 xenopus tro
16	242	8.9	655	2	Q8W0S0_SORBI	Q8w0s0 sorghum bic
17	241	8.9	325	1	B3GT6_MOUSE	Q91z92 mus musculu
18	239.5	8.8	681	2	Q9LV16_ARATH	Q9lv16 arabidopsis
19	237.5	8.8	739	2	Q9SUA8_ARATH	Q9sua8 arabidopsis
20	235	8.7	657	2	Q9LFX6_ARATH	Q9lfx6 arabidopsis
21	233.5	8.6	618	2	Q6AV44_ORYSA	Q6av44 oryza sativ
22	228	8.4	642	2	Q9SSG4_ARATH	Q9ssg4 arabidopsis
23	228	8.4	672	2	Q8RX55_ARATH	Q8rx55 arabidopsis
24	227	8.4	329	1	B3GT6_HUMAN	Q96l58 homo sapien
25	226.5	8.4	385	2	Q7PVV3_ANOGA	Q7pvv3 anopheles g
26	226	8.3	304	2	Q499Z2_HUMAN	Q499z2 homo sapien
27	224.5	8.3	376	2	Q4V698_DROME	Q4v698 drosophila
28	224.5	8.3	382	2	Q9V4W0_DROME	Q9v4w0 drosophila
29	223.5	8.3	108	2	Q4RPY5_TETNG	Q4rpy5 tetraodon n
30	217.5	8.0	651	2	Q84T07_ORYSA	Q84t07 oryza sativ
31	214	7.9	328	2	Q4S321_TETNG	Q4s321 tetraodon n
32	210.5	7.8	447	2	Q5QLP3_ORYSA	Q5qlp3 oryza sativ
33	210.5	7.8	655	2	Q6ZKZ6_ORYSA	Q6zkz6 oryza sativ
34	203	7.5	388	2	Q6DJ37_XENTR	Q6dj37 xenopus tro
35	203	7.5	426	2	Q4S6P1_TETNG	Q4s6p1 tetraodon n
36	202.5	7.5	613	2	Q60VC0_CAEBR	Q60vc0 caenorhabdi
37	201	7.4	421	2	Q640C6_XENLA	Q640c6 xenopus lae
38	200.5	7.4	422	1	B3GT2_HUMAN	O43825 homo sapien
39	200.5	7.4	422	1	B3GT2_PONPY	Q5r5y3 pongo pygma
40	199	7.4	420	2	Q68ET9_XENLA	Q68et9 xenopus lae
41	197.5	7.3	409	2	Q920V2_MUSSI	Q920v2 mus spicile
42	197.5	7.3	422	1	B3GT2_MOUSE	O54905 mus musculu
43	196.5	7.3	422	2	Q3T0R9_BOVIN	Q3t0r9 bos taurus
44	195.5	7.2	330	1	SQV2_CAEEL	Q9n491 caenorhabdi
45	194	7.2	305	2	Q920V5_MUSSI	Q920v5 mus spicile

ALIGNMENTS

RESULT 1
Q8BG28_MOUSE
ID Q8BG28_MOUSE PRELIMINARY; PRT; 504 AA.
AC Q8BG28;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Adult male urinary bladder cDNA, RIKEN full-length enriched library,
DE clone:9530006I10 product:hypothetical Glycosyltransferase family 31
DE containing protein, full insert sequence (Bone marrow macrophage cDNA,
DE RIKEN full-length enriched library, clone:I830149D05
DE product:Hypothetical glycosyltransferase family 31 containing protein,
DE full insert sequence) (Adult male aorta and vein cDNA, RIKEN full-
DE length enriched library, clone:A530062K18 product:hypothetical
DE Glycosyltransferase family 31 containing protein, full insert
DE sequence) (12 days embryo eyeball cDNA, RIKEN full-length enriched
DE library, clone:D230016N13 product:hypothetical Glycosyltransferase
DE family 31 containing protein, full insert sequence) (Beta1,3-N-
DE acetylgalactosaminyltransferase) (11 days pregnant adult female
DE placenta cDNA, RIKEN full-length enriched library, clone:I530025M01
DE product:Hypothetical glycosyltransferase family 31 containing protein,
DE full insert sequence) (Bone marrow macrophage cDNA, RIKEN full-length
DE enriched library, clone:I830033H23 product:Hypothetical
DE glycosyltransferase family 31 containing protein, full insert
DE sequence).
GN Name=B3galnt2; Synonyms=beta3GalNAcT2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impiombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,

RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminiecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavesi G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, and Urinary bladder;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;

```
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN   [10]

  Query Match             100.0%;  Score 2707;  DB 2;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-206;
  Matches  504;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MRNWLVLLCPCVLGAALHLWHLWLRSPPDPHNTGPSAADQSALFPHWKFSHYDVVVGVLS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRNWLVLLCPCVLGAALHLWHLWLRSPPDPHNTGPSAADQSALFPHWKFSHYDVVVGVLS 60

Qy         61 ARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFIIGARGCEVPVEDREDPYSCRLLNITN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFIIGARGCEVPVEDREDPYSCRLLNITN 120

Qy        121 PVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVLYPIVITSLGVFYDASDVGFQRNITVK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVLYPIVITSLGVFYDASDVGFQRNITVK 180

Qy        181 LYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQFILPESFEGTIVWESQDLHGLVSRNL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQFILPESFEGTIVWESQDLHGLVSRNL 240

Qy        241 HRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVAGGFIYTVQEGDALLRSLYSRPQRLAD 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVAGGFIYTVQEGDALLRSLYSRPQRLAD 300

Qy        301 HIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNVPAKLLNFYRWTVESTSFDLLLKTDDD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 HIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNVPAKLLNFYRWTVESTSFDLLLKTDDD 360

Qy        361 CYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVDRTGKWQELEYPSPAYPAFACGSGYVI 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVDRTGKWQELEYPSPAYPAFACGSGYVI 420

Qy        421 SKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGPKRHQDSLWLCEKTCETGMLSSPQYSP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGPKRHQDSLWLCEKTCETGMLSSPQYSP 480

Qy        481 EELSKLWELKELCGDPCQCEAKVR 504
              |||||||||||||||||||||||||
Db        481 EELSKLWELKELCGDPCQCEAKVR 504


RESULT 2
Q5U4F9_MOUSE
ID   Q5U4F9_MOUSE    PRELIMINARY;    PRT;   504 AA.
AC   Q5U4F9;
DT   07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT   07-DEC-2004, sequence version 1.
DT   07-FEB-2006, entry version 10.
DE   UDP-GalNAc:betaGlcNAc beta 1,3-galactosaminyltransferase, polypeptide
DE   2.
GN   Name=B3galnt2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Trophoblast stem cells; STRAIN=B5/EGFP transgenic ICR mice;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B5/EGFP transgenic ICR mice; TISSUE=Trophoblast stem cells;
RG   NIH MGC Project;
RL   Submitted (OCT-2004) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; BC085110; AAH85110.1; -; mRNA.
DR   Ensembl; ENSMUSG00000039242; Mus musculus.
DR   MGI; MGI:2145517; B3galnt2.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008378; F:galactosyltransferase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006486; P:protein amino acid glycosylation; IEA.
DR   InterPro; IPR002659; Glyco_trans_31.
DR   PANTHER; PTHR11214; Glyco_trans_31; 1.
DR   Pfam; PF01762; Galactosyl_T; 1.
KW   Transferase.
SQ   SEQUENCE   504 AA;  57199 MW;  E07A4A1E9B99D76F CRC64;

  Query Match             99.7%;  Score 2698;  DB 2;  Length 504;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-205;
  Matches  502;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MRNWLVLLCPCVLGAALHLWHLWLRSPPDPHNTGPSAADQSALFPHWKFSHYDVVVGVLS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRNWLVLLCPCVLGAALHLWHLWLRSPPDPHNTGPSAADQSALFPHWKFSHYDVVVGVLP 60

Qy         61 ARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFIIGARGCEVPVEDREDPYSCRLLNITN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFIIGARGCEVPVEDREDPYSCRLLNITN 120

Qy        121 PVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVLYPIVITSLGVFYDASDVGFQRNITVK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVLYPIVITSLGVFYDASDVGFQRNITVK 180

Qy        181 LYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQFILPESFEGTIVWESQDLHGLVSRNL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQFILPESFEGTIVWESQDLHGLVSRNL 240

Qy        241 HRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVAGGFIYTVQEGDALLRSLYSRPQRLAD 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVAGGFIYTVQEGDALLRSLYSRPQRLAD 300

Qy        301 HIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNVPAKLLNFYRWTVESTSFDLLLKTDDD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 HIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNVPAKLLNFYRWTVESTSFDLLLKTDDD 360

Qy        361 CYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVDRTGKWQELEYPSPAYPAFACGSGYVI 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVDRTGKWQELEYPGPAYPAFACGSGYVI 420

Qy        421 SKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGPKRHQDSLWLCEKTCETGMLSSPQYSP 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGPKRHQDSLWLCEKTCETGMLSSPQYSP 480

Qy        481 EELSKLWELKELCGDPCQCEAKVR 504
              |||||||||||||||||||||||||
Db        481 EELSKLWELKELCGDPCQCEAKVR 504
```

OM protein - protein search, using sw model

Run on: October 3, 2006, 10:00:07 ; Search time 133.003 Seconds
(without alignments)
3268.776 Million cell updates/sec

Title: US-10-539-450-4_COPY_35_504
Perfect score: 2502
Sequence: 1 PSAADQSALFPHWKFSHYDV.....KLWELKELCGDPCQCEAKVR 470

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2502	100.0	504	2	Q8BG28_MOUSE	Q8bg28 m adult mal
2	2493	99.6	504	2	Q5U4F9_MOUSE	Q5u4f9 mus musculu
3	2236.5	89.4	500	2	Q8NCR0_HUMAN	Q8ncr0 homo sapien
4	2068.5	82.7	427	2	Q59GR3_HUMAN	Q59gr3 homo sapien
5	1776	71.0	488	2	Q5M900_XENTR	Q5m900 xenopus tro
6	1771	70.8	486	2	Q6NRQ1_XENLA	Q6nrq1 xenopus lae
7	1314	52.5	491	2	Q502B3_BRARE	Q502b3 brachydanio
8	1154.5	46.1	326	2	Q5TCI3_HUMAN	Q5tci3 homo sapien
9	1154.5	46.1	382	2	Q96AL7_HUMAN	Q96al7 homo sapien
10	1096.5	43.8	426	2	Q4RPY6_TETNG	Q4rpy6 tetraodon n
11	792	31.7	209	2	Q8BXL0_MOUSE	Q8bxl0 mus musculu
12	606	24.2	155	2	Q4TDH0_TETNG	Q4tdh0 tetraodon n
13	255.5	10.2	663	2	Q6ZKZ3_ORYSA	Q6zkz3 oryza sativ
14	246	9.8	673	2	Q8GXG6_ARATH	Q8gxg6 arabidopsis
15	243.5	9.7	343	2	Q5BL85_XENTR	Q5bl85 xenopus tro
16	242	9.7	655	2	Q8W0S0_SORBI	Q8w0s0 sorghum bic
17	241	9.6	325	1	B3GT6_MOUSE	Q91z92 mus musculu
18	239	9.6	681	2	Q9LV16_ARATH	Q9lv16 arabidopsis
19	237.5	9.5	739	2	Q9SUA8_ARATH	Q9sua8 arabidopsis
20	235	9.4	657	2	Q9LFX6_ARATH	Q9lfx6 arabidopsis
21	233.5	9.3	618	2	Q6AV44_ORYSA	Q6av44 oryza sativ
22	228	9.1	642	2	Q9SSG4_ARATH	Q9ssg4 arabidopsis
23	228	9.1	672	2	Q8RX55_ARATH	Q8rx55 arabidopsis
24	227	9.1	329	1	B3GT6_HUMAN	Q96l58 homo sapien
25	226.5	9.1	385	2	Q7PVV3_ANOGA	Q7pvv3 anopheles g
26	226	9.0	304	2	Q499Z2_HUMAN	Q499z2 homo sapien
27	224.5	9.0	376	2	Q4V698_DROME	Q4v698 drosophila
28	224.5	9.0	382	2	Q9V4W0_DROME	Q9v4w0 drosophila
29	223.5	8.9	108	2	Q4RPY5_TETNG	Q4rpy5 tetraodon n
30	217.5	8.7	651	2	Q84T07_ORYSA	Q84t07 oryza sativ
31	214	8.6	328	2	Q4S321_TETNG	Q4s321 tetraodon n
32	210.5	8.4	447	2	Q5QLP3_ORYSA	Q5qlp3 oryza sativ
33	210.5	8.4	655	2	Q6ZKZ6_ORYSA	Q6zkz6 oryza sativ
34	203	8.1	388	2	Q6DJ37_XENTR	Q6dj37 xenopus tro
35	203	8.1	426	2	Q4S6P1_TETNG	Q4s6p1 tetraodon n
36	202.5	8.1	613	2	Q60VC0_CAEBR	Q60vc0 caenorhabdi
37	201	8.0	421	2	Q640C6_XENLA	Q640c6 xenopus lae
38	200.5	8.0	422	1	B3GT2_HUMAN	O43825 homo sapien
39	200.5	8.0	422	1	B3GT2_PONPY	Q5r5y3 pongo pygma
40	199	8.0	420	2	Q68ET9_XENLA	Q68et9 xenopus lae
41	197.5	7.9	409	2	Q920V2_MUSSI	Q920v2 mus spicile
42	197.5	7.9	422	1	B3GT2_MOUSE	O54905 mus musculu
43	196.5	7.9	422	2	Q3T0R9_BOVIN	Q3t0r9 bos taurus
44	195.5	7.8	330	1	SQV2_CAEEL	Q9n491 caenorhabdi
45	194	7.8	305	2	Q920V5_MUSSI	Q920v5 mus spicile

ALIGNMENTS

RESULT 1
Q8BG28_MOUSE
ID Q8BG28_MOUSE PRELIMINARY; PRT; 504 AA.
AC Q8BG28;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Adult male urinary bladder cDNA, RIKEN full-length enriched library,
DE clone:9530006I10 product:hypothetical Glycosyltransferase family 31
DE containing protein, full insert sequence (Bone marrow macrophage cDNA,
DE RIKEN full-length enriched library, clone:I830149D05
DE product:Hypothetical glycosyltransferase family 31 containing protein,
DE full insert sequence) (Adult male aorta and vein cDNA, RIKEN full-
DE length enriched library, clone:A530062K18 product:hypothetical
DE Glycosyltransferase family 31 containing protein, full insert
DE sequence) (12 days embryo eyeball cDNA, RIKEN full-length enriched
DE library, clone:D230016N13 product:hypothetical Glycosyltransferase
DE family 31 containing protein, full insert sequence) (Beta1,3-N-
DE acetylgalactosaminyltransferase) (11 days pregnant adult female
DE placenta cDNA, RIKEN full-length enriched library, clone:I530025M01
DE product:Hypothetical glycosyltransferase family 31 containing protein,
DE full insert sequence) (Bone marrow macrophage cDNA, RIKEN full-length
DE enriched library, clone:I830033H23 product:Hypothetical
DE glycosyltransferase family 31 containing protein, full insert
DE sequence).
GN Name=B3galnt2; Synonyms=beta3GalNAcT2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impiombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,

RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminiecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavesi G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Bone marrow, Eyeball,
RC Placenta, and Urinary bladder;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, and Urinary bladder;
RA Adachi J., Arakawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;

```
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN   [10]

  Query Match          100.0%;  Score 2502;  DB 2;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-191;
  Matches  470;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PSAADQSALFPHWKFSHYDVVVGVLSARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 PSAADQSALFPHWKFSHYDVVVGVLSARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFI 94

Qy         61 IGARGCEVPVEDREDPYSCRLLNITNPVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 IGARGCEVPVEDREDPYSCRLLNITNPVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVL 154

Qy        121 YPIVITSLGVFYDASDVGFQRNITVKLYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 YPIVITSLGVFYDASDVGFQRNITVKLYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQ 214

Qy        181 FILPESFEGTIVWESQDLHGLVSRNLHRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 FILPESFEGTIVWESQDLHGLVSRNLHRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVA 274

Qy        241 GGFIYTVQEGDALLRSLYSRPQRLADHIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 GGFIYTVQEGDALLRSLYSRPQRLADHIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNV 334

Qy        301 PAKLLNFYRWTVESTSFDLLLKTDDDCYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 PAKLLNFYRWTVESTSFDLLLKTDDDCYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVD 394

Qy        361 RTGKWQELEYPSPAYPAFACGSGYVISKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 RTGKWQELEYPSPAYPAFACGSGYVISKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGP 454

Qy        421 KRHQDSLWLCEKTCETGMLSSPQYSPEELSKLWELKELCGDPCQCEAKVR 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 KRHQDSLWLCEKTCETGMLSSPQYSPEELSKLWELKELCGDPCQCEAKVR 504


RESULT 2
Q5U4F9_MOUSE
ID   Q5U4F9_MOUSE    PRELIMINARY;    PRT;    504 AA.
AC   Q5U4F9;
DT   07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT   07-DEC-2004, sequence version 1.
DT   07-FEB-2006, entry version 10.
DE   UDP-GalNAc:betaGlcNAc beta 1,3-galactosaminyltransferase, polypeptide
DE   2.
GN   Name=B3galnt2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Trophoblast stem cells; STRAIN=B5/EGFP transgenic ICR mice;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B5/EGFP transgenic ICR mice; TISSUE=Trophoblast stem cells;
RG   NIH_MGC Project;
RL   Submitted (OCT-2004) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---------------------------------------------------------------------
DR   EMBL; BC085110; AAH85110.1; -; mRNA.
DR   Ensembl; ENSMUSG00000039242; Mus musculus.
DR   MGI; MGI:2149517; B3galnt2.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008378; F:galactosyltransferase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006486; P:protein amino acid glycosylation; IEA.
DR   InterPro; IPR002659; Glyco_trans_31.
DR   PANTHER; PTHR11214; Glyco_trans_31; 1.
DR   Pfam; PF01762; Galactosyl_T; 1.
KW   Transferase.
SQ   SEQUENCE   504 AA;  57199 MW;  E07A4A1E9B99D76F CRC64;

  Query Match          99.6%;  Score 2493;  DB 2;  Length 504;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-190;
  Matches  468;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PSAADQSALFPHWKFSHYDVVVGVLSARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 PSAADQSALFPHWKFSHYDVVVGVLSARNNHELRNVIRNTWLKNLLHHPTLSQRVLVKFI 94

Qy         61 IGARGCEVPVEDREDPYSCRLLNITNPVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 IGARGCEVPVEDREDPYSCRLLNITNPVLNQEIEAFSFPEDASSSRLSEDRVVSVSFRVL 154

Qy        121 YPIVITSLGVFYDASDVGFQRNITVKLYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 YPIVITSLGVFYDASDVGFQRNITVKLYQTEQEEALFIARFSPPSCGVQVNKLWYKPVEQ 214

Qy        181 FILPESFEGTIVWESQDLHGLVSRNLHRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 FILPESFEGTIVWESQDLHGLVSRNLHRVTVNDGGGVLRVLAAGEGALPHEFMEGVEGVA 274

Qy        241 GGFIYTVQEGDALLRSLYSRPQRLADHIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 GGFIYTVQEGDALLRSLYSRPQRLADHIQDLQVEDALLQEESSVHDDIVFVDVVDTYRNV 334

Qy        301 PAKLLNFYRWTVESTSFDLLLKTDDDCYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVD 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 PAKLLNFYRWTVESTSFDLLLKTDDDCYIDLEAVFNRIAQKNLDGPNFWWGNFRLNWAVD 394

Qy        361 RTGKWQELEYPSPAYPAFACGSGYVISKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGP 420
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 RTGKWQELEYPGPAYPAFACGSGYVISKDIVDWLAGNSRRLKTYQGEDVSMGIWMAAIGP 454

Qy        421 KRHQDSLWLCEKTCETGMLSSPQYSPEELSKLWELKELCGDPCQCEAKVR 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 KRHQDSLWLCEKTCETGMLSSPQYSPEELSKLWELKELCGDPCQCEAKVR 504


RESULT 3
Q8NCR0_HUMAN
ID   Q8NCR0_HUMAN    PRELIMINARY;    PRT;    500 AA.
AC   Q8NCR0;
DT   01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT   01-OCT-2002, sequence version 1.
DT   07-FEB-2006, entry version 17.
DE   UDP-GalNAc:betaGlcNAc beta 1,3-galactosaminyltransferase, polypeptide
```